Dzero SAM and SAM-Grid plans

Lee Lueking
OATF
October 2, 2002

Overview

- The Ultimate Goal: SAM-Grid
- Where we are right now and how it plays w/ Regional Analysis Centers
- The plan for reaching the Holy Grail
- Timeline

SAM and the Grid

SAM-Grid

What is SAM-Grid?

- Project to include Job and Information Management (JIM) with the SAM Data Management System
- Project started in 2001 as part of the PPDG collaboration to handle D0's expanded needs.
- Current SAM-Grid team includes:
  - Andrew Baranovski, Gabriele Garzoglio, Lee Lueking, Siddharth Patil, Abhishek Rana, Dane Skow, Igor Terekhov, Rod Walker (Imperial College), Jae Yu (U. Texas Arlington)
  - Collaboration with U. Wisconsin Condor team.
- **http://www-d0.fnal.gov/computing/grid**

The Goal

- Enable fully distributed computing for the DZero (and CDF), by enhancing SAM and incorporating standard Grid tools and protocols. Developing new solutions for Grid computing in a secure and accountable environment.
- The SAM grid-ification is funded by PPDG and GridPP. The collaborators we are working with include the Condor Team (via PPDG) and Imperial College (via GridPP)
- We are communicating with other groups working on Grid technologies as well (EDG among them).
- Regular CDF/DZero joint grid meetings
- We promote interoperability and code reuse

Condor
High Throughput Computing

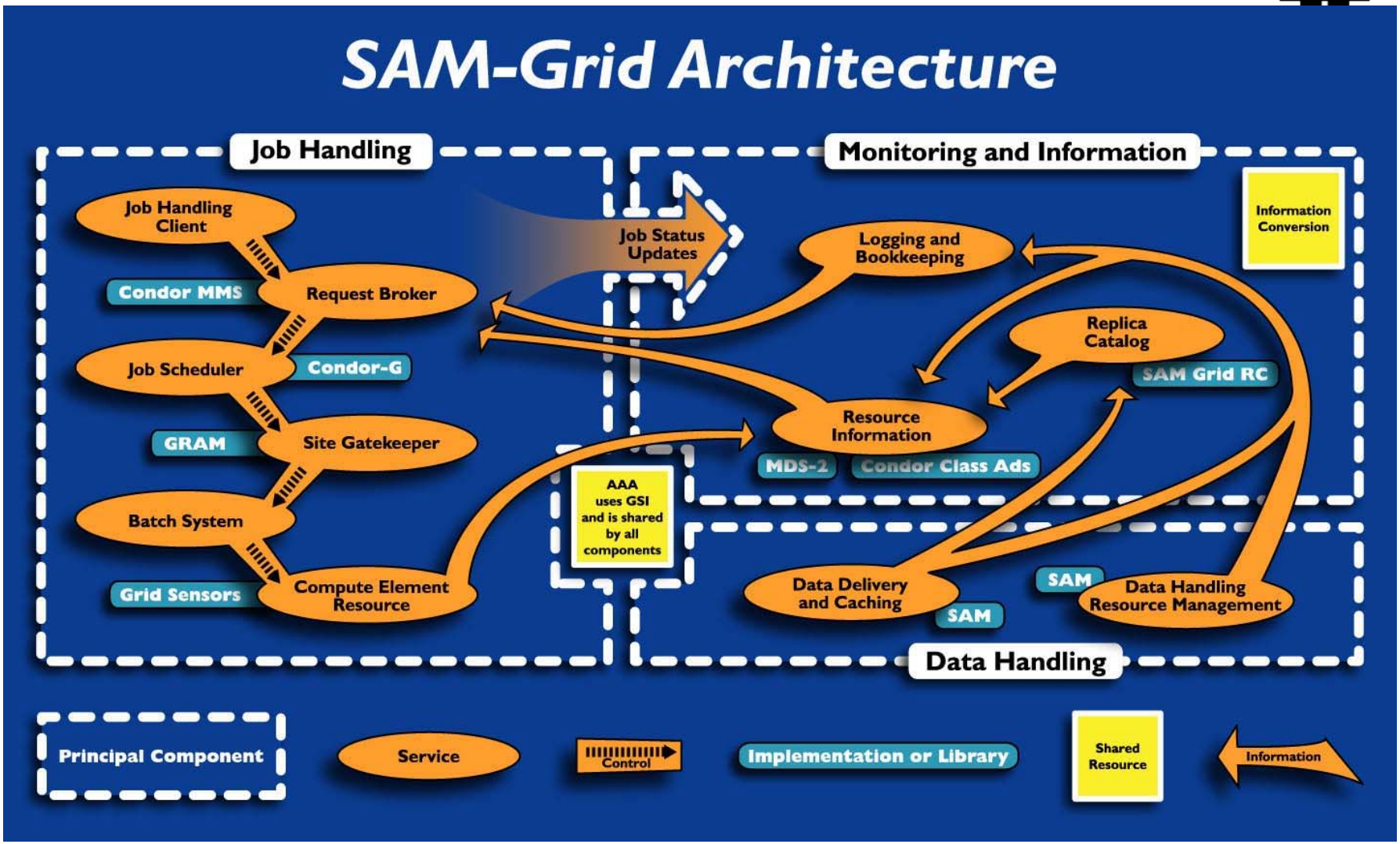
the globus project™

CDF II

PPDG

Major Components

- Job Definition and Management: The preliminary job management architecture is aggressively based on the Condor technology provided by through our collaboration with University of Wisconsin CS Group.
- Monitoring and Information Services: We assign a critical role to this part of the system and widen the boundaries of this component to include all services that provide, or receive, information relevant for job and data management.
- Data Handling: The existing SAM Data Handling system, when properly abstracted, plays a principal role in the overall architecture and has direct effects on the Job Management services.

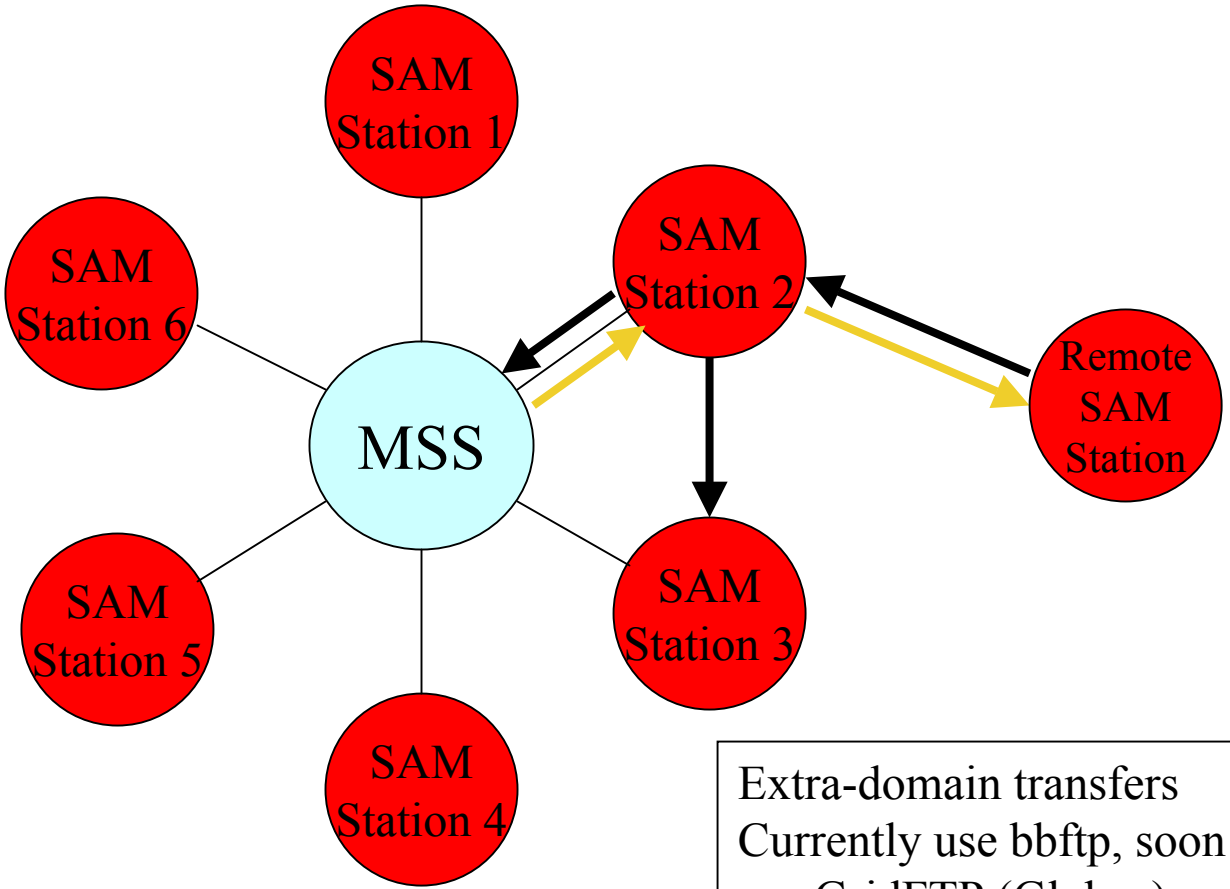

SAM-Grid Architecture
Job Handling
Job Handling Client
Condor MMS
Request Broker
Job Scheduler
Condor-G
GRAM
Site Gatekeeper
Batch System
Grid Sensors
Compute Element Resource
Job Status Updates
AAA uses GSI and is shared by all components
Monitoring and Information
Logging and Bookkeeping
Information Conversion
Replica Catalog
SAM Grid RC
Resource Information
MDS-2
Condor Class Ads
Data Delivery and Caching
SAM
SAM
Data Handling Resource Management
Data Handling
Principal Component
Service
Control
Implementation or Library
Shared Resource
Information

SAM as it is now

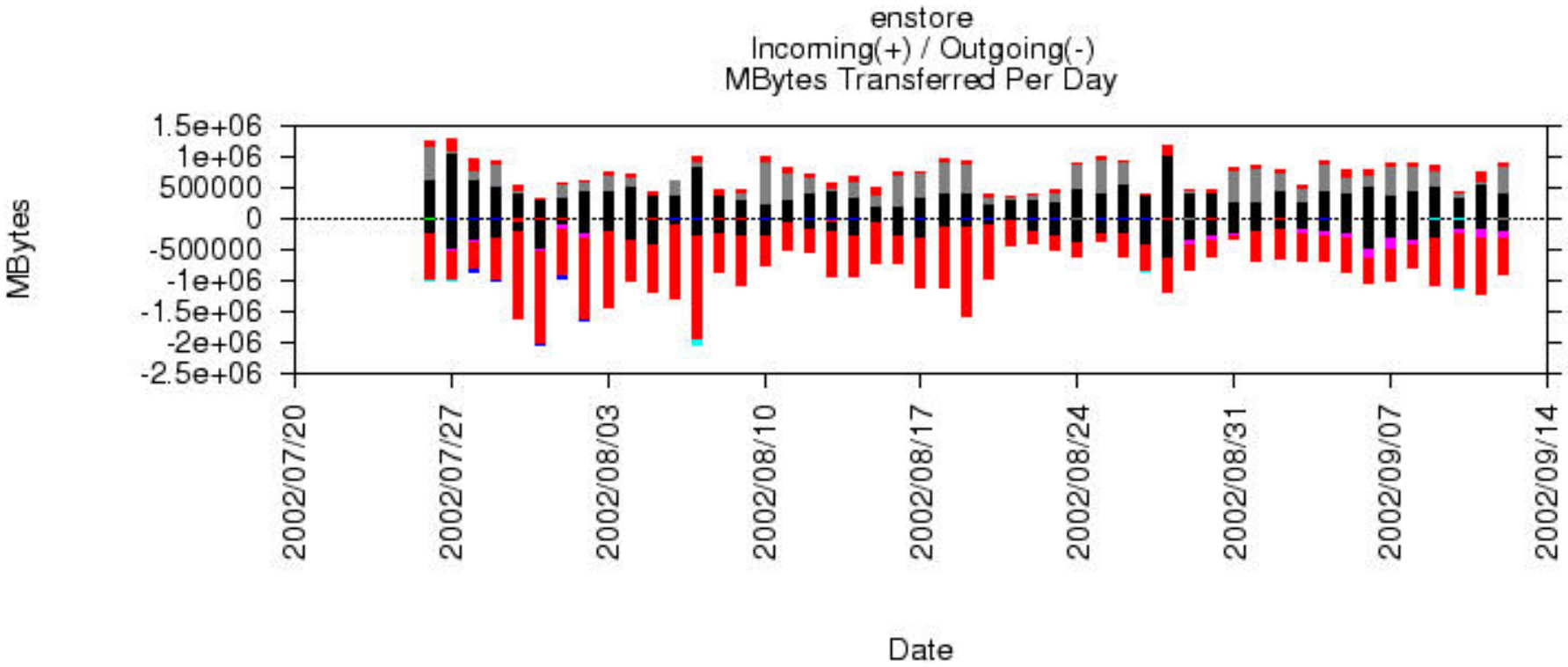
SAM-Now

Data to and from Remote Sites

Station Configuration
- Replica location
  - Prefer
  - Avoid
- Forwarding
  - File stores can be forwarded through other stations
- Routing
  - Routes for file transfers are configurable
- Remote Staging
  - Cache & SAM stagers are maintained on D0mino for remote enstore access

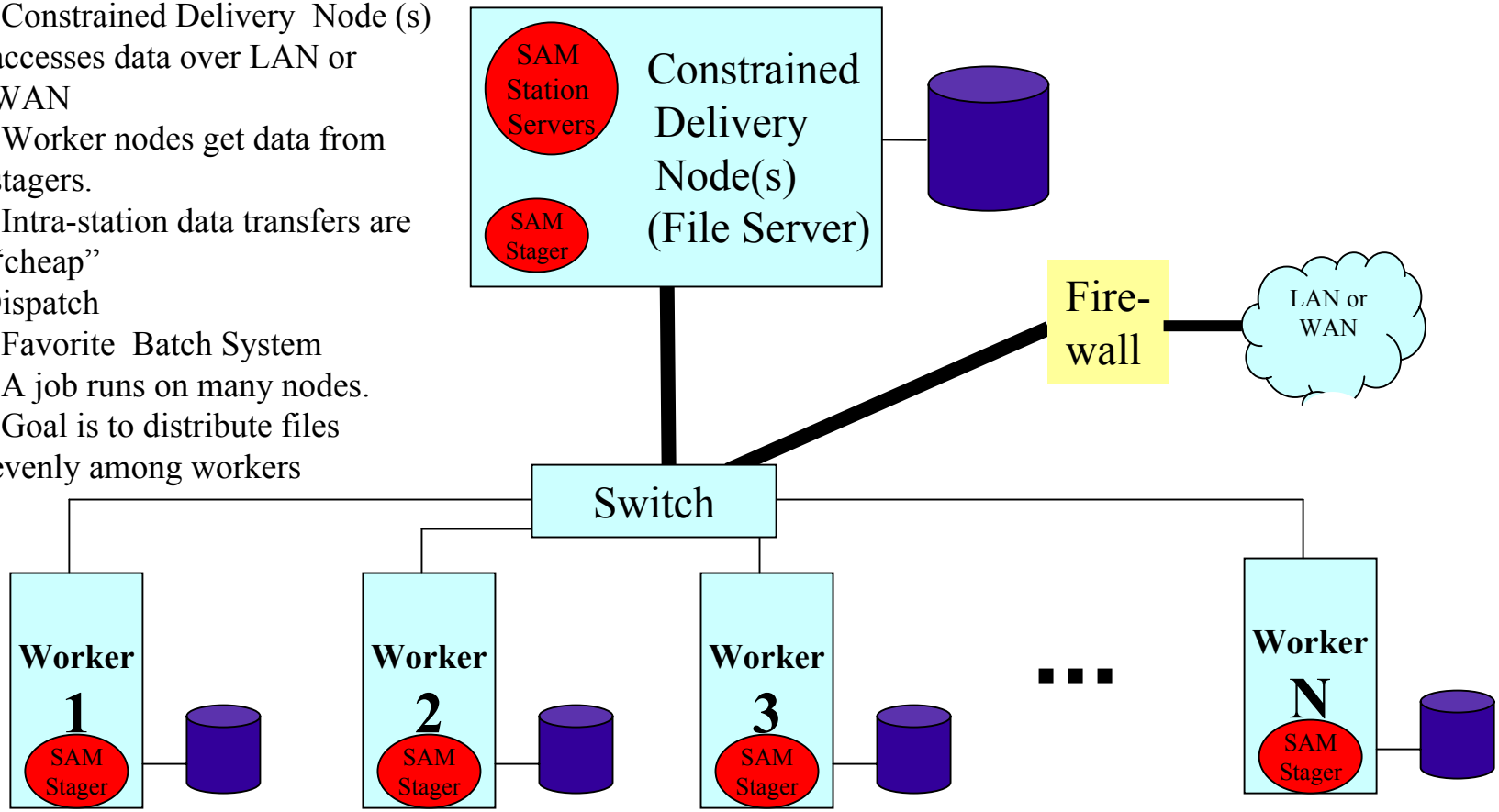


Extra-domain transfers
Currently use bbftp, soon use GridFTP (Globus) (parallel transfer protocols)


enstore
Incoming(+) / Outgoing(-)
MBytes Transferred Per Day
MBytes
1.5e+06
1e+06
500000
0
-500000
-1e+06
-1.5e+06
-2e+06
-2.5e+06
2002/07/20
2002/07/27
2002/08/03
2002/08/10
2002/08/17
2002/08/24
2002/08/31
2002/09/07
2002/09/14
Date
Stations:
D0Mainz
cab
central-analysis
cinvestav-station
d0karlsruhe
datalogger-d0olb
datalogger-d0olc
fnal-farm
hoeve
imperial-test
princeton-d0
tata-d0-mcfarm
triviaal
uta-analysis
uta-d0-grid
uta-hep
wuppertal

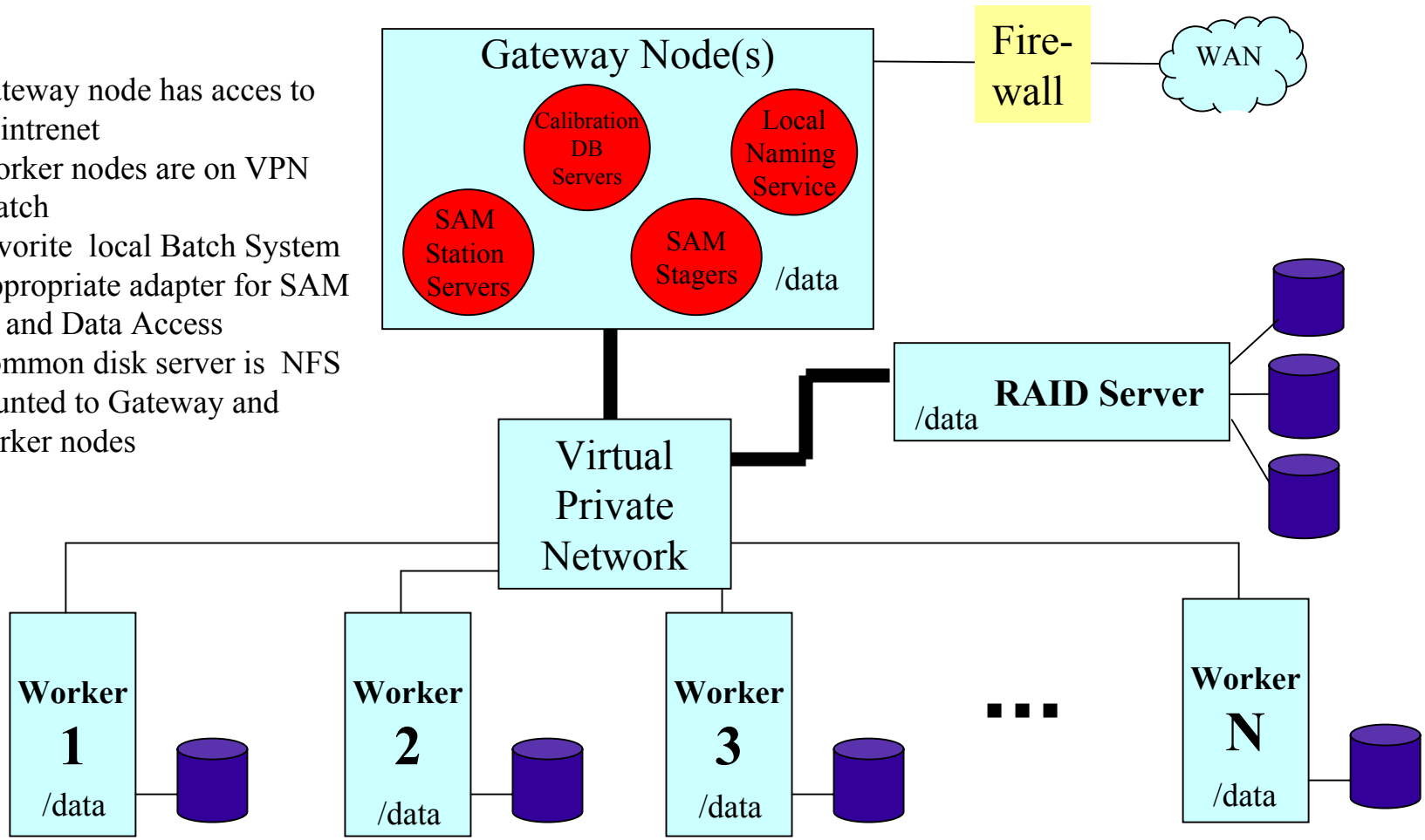
SAM Station: Dzero Distributed Cache Farm on Public Network

- Network
  - Constrained Delivery Node (s) accesses data over LAN or WAN
  - Worker nodes get data from stagers.
  - Intra-station data transfers are "cheap"
- Job Dispatch
  - Favorite Batch System
  - A job runs on many nodes.
  - Goal is to distribute files evenly among workers



SAM Station: Shared Cache Configuration w/ VPN

- Network
  - Gateway node has acces to the intrenet
  - Worker nodes are on VPN
- Job Dispatch
  - Favorite local Batch System
  - Appropriate adapter for SAM
- Software and Data Access
  - Common disk server is NFS mounted to Gateway and Worker nodes



More Details

- If the gateway node is behind a firewall, ports > 1024 need to be unrestricted to d0ora1.fnal.gov, d0ora3.fnal.gov, and d0mino.fnal.gov.
- Gateway runs sam servers, special setup, user sam account.
- Runs bbftp or GridFTP demon for parallel transfers.
- SAM is distributed to clients via fnal ups/upd products distribution and versioning.
- SAM shared cache configuration has been successfully implemented at Karlsruhe. Many other sites are interested in using it.
- Anticipate NFS and RAID server bottlenecks.
- Calibration DB servers are caching proxies connected through primary servers at FNAL to the central data base.

How we get to SAM-Grid

Quest for the Grail

The steps in getting to SAM-Grid

- JIM Project
  - Job Management
  - Job Description Language
  - Information Service
  - Testbed deployment at selected sites. Includes 1) GRAM (Globus Resource Allocation Manager) gatekeeper and use of local scheduler, 2) MDS (Monitoring and Discovery Service)
  - Prototype is now available (Igor will say more)
- Grid Security (AAA) using GSI (or other).
  - Have GridFTP working as a sam transfer protocal
  - Latest bbftp also has GSI security plug-in feature
  - Need VO and User-level certificate authentication and authorization.
  - Job submission to authorized compute resources

The steps in getting to SAM-Grid

- Uniform process submission interface (sam submit)
- dCache integration for rate adapting and remote station file serving.
- Understand the modularization of the data handling and storage interfaces
- Generalized HSM Adapters to employ:
  - HPSS @ Lyons, enstore @ Lancaster, or other MSS.
  - Network attached files (file url)
  - SRM interface
  - Additional dCache features
  - Other Storage Elements like disk farm
- D0 Run Time Environment will allow running on resources not tailored to D0 (no D0 installation).
- Site Autonomous SAM station and site resource management
- One-step SAM installation and registration (In the long term)

Possible Timeline